CLAIMS

We claim:

B1

1. A software implemented methodology for determining the pK_a of a molecule of interest comprising the following steps:

a) determine the hierarchical atom type connectivity tree noting the number of times each atom type occurs for each ionizable group on each molecule of a series of molecules for which an experimentally determined pK_a is known;

b) place each of the determined atom types noting the number of times each atom type occurs for each hierarchically determined connectivity tree in a row of a data table along with the experimentally determined pK_a of the molecule from which the tree was determined;

c) using the partial least squares (PLS) statistical methodology, extract coefficients associated with each atom type represented at each hierarchical level;

d) determine the hierarchical atom type connectivity tree noting the number of times each atom type occurs for the molecule of interest; and

e) multiply the number of occurrences of each atom type in the molecule of interest by the PLS coefficient determined for that atom type and sum the resulting multiplications to obtain the predicted pK_a.

2. The method of claim 1 in which each atom type from the hierarchical atom connectivity tree in steps a and d for each ionizable group on each molecule is placed into a separate bin in a bit string.

3. The method of claim 2 in which the extracted coefficients are associated with the

appropriate bin in the bit string.

4.(new) The method of claim 1 in which in step 1.a the ionizable group is selected from the group consisting of C.3 - sp^3 carbon, N.3 - sp^3 nitrogen, N.2 - sp^2 nitrogen, N.ar - aromatic nitrogen, N.am - amide nitrogen, N.pl3 - planar sp^3 nitrogen, O.3 - sp^3 oxygen, and S.3 - sp^3 sulfur.

5.(new) The method of claim 4 in which each atom type from the hierarchical atom connectivity tree in steps a and d for each ionizable group on each molecule is placed into a separate bin in a bit string.

6.(new) The method of claim 5 in which the extracted coefficients are associated with the appropriate bin in the bit string.

7.(new) A software implemented methodology for predicting the pK_a of a molecule of interest using atom types and partial least squares (PLS) comprising the following steps:

a) determine the hierarchical atom type connectivity tree noting the number of times each atom type occurs for each ionizable atom type on each molecule of a series of molecules for which an experimentally determined pK_a is known;

b) place each of the determined atom types noting the number of times each atom type occurs for each hierarchically determined connectivity tree in a row of a data table along with the experimentally determined pK_a of the molecule from which the tree was determined;

c) using the partial least squares (PLS) statistical methodology, extract coefficients associated with each atom type represented at each hierarchical level;

d) determine the hierarchical atom type connectivity tree noting the number of times

B1

each atom type occurs for each ionizable atom type for the molecule of interest; and

e) multiply the number of occurrences of each atom type in the molecule of interest by the PLS coefficient determined for that atom type and sum the resulting multiplications to obtain the predicted pK_a.

8.(new) The method of claim 7 in which each atom type from the hierarchical atom connectivity tree in steps a and d for each ionizable atom type on each molecule is placed into a separate bin in a bit string.

9.(new) The method of claim 8 in which the extracted coefficients are associated with the appropriate bin in the bit string.

10.(new) The method of claim 7 in which in step 1.a the ionizable atom type is selected from the group consisting of C.3 - sp^3 carbon, N.3 - sp^3 nitrogen, N.2 - sp^2 nitrogen, N.ar - aromatic nitrogen, N.am - amide nitrogen, N.pl3 - planar sp^3 nitrogen, O.3 - sp^3 oxygen, and S.3 - sp^3 sulfur.

11.(new) The method of claim 10 in which each atom type from the hierarchical atom connectivity tree in steps a and d for each ionizable atom type on each molecule is placed into a separate bin in a bit string.

12.(new) The method of claim 11 in which the extracted coefficients are associated with the appropriate bin in the bit string.

B1